

## Claims

1. A method for optimizing data transmission in a data transfer system comprising:

5 monitoring a level of data transfer congestion within the data transfer system, the monitoring including marking data during data transfer congestion and detecting marked data; and

adjusting a data transfer rate corresponding to the level of data transfer congestion;

10 wherein the adjusting includes reducing the data transfer rate in direct correlation to the level of data transfer congestion as indicated by each marked data and increasing the data transfer rate in direct correlation to a lack of data transfer congestion as indicated by unmarked data during a round trip time (RTT).

15 2. A method for optimizing data transmission in a data transfer system as recited in claim 1, wherein the marked data is a data packet that is marked with data congestion information.

20 3. A method for optimizing data transmission in a data transfer system as recited in claim 1, wherein the unmarked data is a data packet without data congestion information.

25 4. A method for optimizing data transmission in a data transfer system as recited in claim 1, wherein the marking data during data transfer congestion includes:

sending a data packet to a routing mechanism;

determining a fraction of the input buffer of the routing mechanism that is filled;

randomly marking the data packet according to a probability identical to the fraction of the input buffer that is filled, the random marking indicating data transfer congestion; and

5 generating an acknowledgement data by a recipient of the data packet, the acknowledgment data being marked if the data packet is marked.

5. A method for optimizing data transmission in a data transfer system as recited in claim 4, wherein the detecting marked data includes:

10 receiving the acknowledgment data from the recipient of the data packet; and

analyzing the acknowledgement data to determine if the data packet was marked.

6. A method for optimizing data transmission in a data transfer system as recited in claim 2, wherein the reducing further includes decreasing the data transmission  
15 rate by one data packet per round trip time (PRTT) for every marked packet detected.

7. A method for optimizing data transmission in a data transfer system as recited in claim 3, wherein the increasing further includes transmitting one additional data packet per round trip time (PRTT) where only the unmarked data packets are detected  
20 during a previous round trip time.

8. A method for optimizing data transmission in a data transfer system as recited in claim 2, wherein the marking the data packet includes setting data congestion information in an internet protocol header of the data packet.

9. A method for optimizing data transmission in a data transfer system as recited in claim 2, wherein the generating acknowledgment data includes setting data congestion information in an acknowledgment header if the data packet is marked.

10. A method for optimizing data transmission in a data transfer system as recited in claim 4, wherein the acknowledgement data is a positive acknowledgement (ACK).

11. A method for optimizing data transmission in a data transfer system as recited in claim 1, wherein the data transfer system includes at least a sending host, a sending switch, and a data recipient.

12. A method for optimizing data transmission in a data transfer system as recited in claim 11, wherein the data recipient includes at least one of a receiving host and a receiving switch connected to the receiving host.

13. A network system for actively controlling congestion to optimize throughput, comprising:

a sending host being configured to send packet traffic at a set rate;

a sending switch for receiving the packet traffic, the sending switch including,

a input buffer for receiving the packet traffic at the set rate, the input buffer being actively monitored to ascertain a capacity level;

code for setting a probability factor that is correlated to the capacity level,  
5 the probability factor increasing as the capacity level increases and decreasing as the capacity level decreases;

code for randomly generating a value, the value being indicative of whether packets being sent by the sending switch are to be marked with a congestion indicator; and

10 transmit code forwarding the packet traffic out of the sending switch, the packet traffic including one of marked packets and unmarked packets; and

a receiving end being configured as the recipient of the packet traffic and configured to generate acknowledgment packets back to the sending host, the acknowledgment packets being marked with the congestion indicator when receiving  
15 marked packets and not being marked with the congestion indicator when receiving unmarked packets.

14. A network system for actively controlling congestion to optimize throughput as recited in claim 13, wherein the sending host is configured to monitor the  
20 acknowledgment packets and to adjust the set rate based on whether the acknowledgment packets are marked with the congestion indicator.

15. A network system for actively controlling congestion to optimize throughput as recited in claim 13, wherein the set rate is a number of packets sent per round trip time (RTT) as determined by a congestion window in the sending host.

5 16. A network system for actively controlling congestion to optimize throughput as recited in claim 15, wherein the congestion window operates to limit the amount of data that can be transmitted by the sending host before the acknowledgement packet is received.

10 17. A network system for actively controlling congestion to optimize throughput as recited in claim 13, wherein the sending host decreases the set rate every time one of the marked packets is detected.

15 18. A network system for actively controlling congestion to optimize throughput as recited in claim 17, wherein the set rate is decreased by one packet per round trip time (PRTT) for each of the marked packets that is detected by the sending host.

20 19. A network system for actively controlling congestion to optimize throughput as recited in claim 17, wherein the sending host increases the set rate when no marked packets are detected per round trip time (PRTT).

20. A network system for actively controlling congestion to optimize throughput as recited in claim 19, wherein the set rate is increased by one packet per round trip time (PRTT).

5 21. A network system for actively controlling congestion to optimize throughput as recited in claim 13, wherein the actively monitored is an examining of the input buffer and a determining of an amount of the packets in the input buffer.

10 22. A network system for actively controlling congestion to optimize throughput as recited in claim 13, wherein the capacity level is a fraction of the input buffer that is filled with the packets.

15 23. A network system for actively controlling congestion to optimize throughput as recited in claim 13, wherein the probability factor is a percentage probability that one of the packets sent from the sending switch to the receiving end will be marked.

20 24. A network system for actively controlling congestion to optimize throughput as recited in claim 13, wherein the value is a randomly generated number between 1 and 100.

25. A network system for actively controlling congestion to optimize throughput as recited in claim 13, wherein the congestion indicator is data in the IP header of the packets showing that data transfer congestion exists.

5 26. A method for actively controlling congestion to optimize throughput comprising:

transferring a data packet to an input buffer of a sending switch at a set rate;

monitoring the input buffer to ascertain a capacity level of the input buffer;

10 setting a probability factor that is correlated to the capacity level of the input buffer, the probability factor increasing as the capacity level increases and decreasing as the capacity level decreases;

randomly generating a value, the value being indicative of whether the data packet sent by the sending switch is to be marked with a congestion indicator;

15 forwarding the data packet out of the sending switch to a recipient, the data packet being one of a marked data packet and an unmarked data packet; and

generating an acknowledgment packet to be sent from the recipient to the sending host, the acknowledgment packet being marked with the congestion indicator when receiving the marked data packet and not being marked with the congestion indicator when receiving the unmarked data packet.

20 27. A method for actively controlling congestion to optimize throughput as recited in claim 26, further comprising:

monitoring the acknowledgment packets; and

adjusting the set rate based on whether the acknowledgment packet is marked with the congestion indicator.

28. A method for actively controlling congestion to optimize throughput as recited in claim 27, wherein the set rate is a number of data packets sent per round trip time (PRTT) as determined by a congestion window in the sending host.

29. A method for actively controlling congestion to optimize throughput as recited in claim 28, wherein the congestion window operates to limit the amount of the data packets that can be transmitted by the sending host before the acknowledgement packet is received.

30. A method for actively controlling congestion to optimize throughput as recited in claim 26, wherein the sending host decreases the set rate every time one of the marked data packets is detected.

31. A method for actively controlling congestion to optimize throughput as recited in claim 30, wherein the set rate is decreased by one data packet per round trip time (PRTT) for each of the marked data packet that is detected by the sending host.

32. A method for actively controlling congestion to optimize throughput as recited in claim 30, wherein the sending host increases the set rate when no marked data packets are detected per round trip time (PRTT)



33. A method for actively controlling congestion to optimize throughput as recited in claim 32, wherein the set rate is increased by one data packet per round trip time (PRTT).

5

34. A method for actively controlling congestion to optimize throughput as recited in claim 26, wherein the actively monitored is an examining of the input buffer and a determining of an amount of the packets in the input buffer.

10 35. A method for actively controlling congestion to optimize throughput as recited in claim 26, wherein the capacity level is a fraction of the input buffer that is filled with a plurality of the data packet.

15 36. A method for actively controlling congestion to optimize throughput as recited in claim 26, wherein the probability factor is a percentage probability that the data packet sent from the sending switch to the receiving end will be marked.

20 37. A network system for actively controlling congestion to optimize throughput as recited in claim 26, wherein the value is a randomly generated number between 1 and 100.

38. A network system for actively controlling congestion to optimize throughput as recited in claim 26, wherein the congestion indicator is data in the IP header of the data packet showing that data congestion exists.

5 39. A method for actively controlling congestion to optimize throughput comprising:

transferring a data packet to an input buffer of a sending switch at a set rate;

monitoring the input buffer to ascertain a capacity level of the input buffer;

10 setting a probability factor that is correlated to the capacity level of the input buffer, the probability factor increasing as the capacity level increases and decreasing as the capacity level decreases;

randomly generating a value, the value being indicative of whether the data packet sent by the sending switch is to be marked with a congestion indicator;

15 forwarding the data packet out of the sending switch to a recipient, the data packet being one of a marked data packet and an unmarked data packet;

generating an acknowledgment packet to be sent from the recipient to the sending host, the acknowledgment packet being marked with the congestion indicator when receiving the marked data packet and not being marked with the congestion indicator when receiving the unmarked data packet;

20 monitoring the acknowledgment packets; and

adjusting the set rate based on whether the acknowledgment packet is marked with the congestion indicator.

